THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXV JANUARY 1956 NUMBER 1

Diffused Emitter and Base Silicon Transistors*

By M. TANENBAUM and D. E. THOMAS

(Manuscript received October 21, 1955)

Silicon n-p-n transistors have been made in which the base and emitter regions were produced by diffusing impurities from the vapor phase. Transistors with base layers 3.8 × 10^{-4}-cm thick have been made. The diffusion techniques and the processes for making electrical contact to the structures are described.

The electrical characteristics of a transistor with a maximum alpha of 0.97 and an alpha-cutoff of 120 mc/sec are presented. The manner in which some of the electrical parameters are determined by the distribution of the doping impurities is discussed. Design data for the diffused emitter, diffused base structure is calculated and compared with the measured characteristics.

INTRODUCTION

The necessity of thin base layers for high-frequency operation of transistors has long been apparent. One of the most appealing techniques for controlling the distribution of impurities in a semiconductor is the diffusion of the impurity into the solid semiconductor. The diffusion coefficients of Group III acceptors and Group V donors into germanium and silicon are sufficiently low at judiciously selected temperatures so

* A portion of the material of this paper was presented at the Semiconductor Device Conference of the Institute of Radio Engineers, Philadelphia, Pa., June, 1955.

that it is possible to envision transistors with base layer thicknesses of a micron and frequency response of several thousand megacycles per second.

A major deterent to the application of diffusion to silicon transistor fabrication in the past was the drastic decrease in lifetime which generally occurs when silicon is heated to the high temperatures required for diffusion. There was also insufficient knowledge of the diffusion parameters to permit the preparation of structures with controlled layer thicknesses and desired dopings. Recently the investigations of C. S. Fuller and coworkers have produced detailed information concerning the diffusion of Group III and Group V elements in silicon. This information has made possible the controlled fabrication of transistors with base layers sufficiently thin that high alphas are obtained even though the lifetime has been reduced to a fraction of a microsecond. In a cooperative program with Fuller, diffusion structures were produced which have permitted the fabrication of transistors whose electrical behavior closely approximates the behavior anticipated from the design. This paper describes these techniques which have resulted in high alpha silicon transistors with alpha-cutoff of over 100 mc/sec.

1.0 FABRICATION OF THE TRANSISTORS

Fuller's work[1] has shown that in silicon the diffusion coefficient of a Group III acceptor is usually 10 to 100 times larger than that of the Group V donor in the same row in the periodic table at the same temperatures. These experiments were performed in evacuated silica tubes using the Group III and Group V elements as the source of diffusant. Under these conditions a particular steady state surface concentration of the diffusant is produced and the depth of diffusion is sensitive to this concentration as well as to the diffusion coefficient. The experiments show that the effective steady state surface concentration of the donor impurities produced under these conditions is ten to one hundred times greater than that of the acceptor impurities. Thus, by the simultaneous diffusion of selected donor and acceptor impurities into n-type silicon an n-p-n structure will result. The first n-layer forms because the surface concentration of the donor is greater than that of the acceptor. The p-layer is produced because the acceptor diffuses faster than the donor and gets ahead of it. The final n-region is simply the original background doping of the n-type silicon sample. It has been possible to produce n-p-n structures by the simultaneous diffusion of several combinations of donors and acceptors. Often, however, the diffusion coefficients and surface concentrations of the donors and acceptors are such that opti-

[1] C. S. Fuller, private communication.

mum layer thicknesses (see Sections 3 and 4) are not produced by simultaneous diffusion. In this case, one of the impurities is started ahead of the other in a prior diffusion, and then the other impurity is diffused in a second operation.

With the proper choice of diffusion temperatures and times it has been possible to make n-p-n structures with base layer thicknesses of 2×10^{-4} cm. The uniformity of the layers in a given specimen is better than ten per cent of the layer thickness. Fig. 1 illustrates the uniformity of the layers. This figure is an enlarged photograph of a view perpendicular to the surface of the specimen. A bevel which makes an angle of five degrees with the original surface has been polished on the specimen. This angle magnifies the layer thickness by 11.5. The layer is defined by an etchant which preferentially stains p-type silicon[1] and the width of the layer is measured with a calibrated microscope.

After diffusion the entire surface of the silicon wafer is covered with the diffused n- and p-type layers, see Fig. 2(a). Electrical contact must now be made to the three regions of the device. The base contact can be made by polishing a bevel on the specimen to expose and magnify the base layer and then alloying a lead to this region by the same tech-

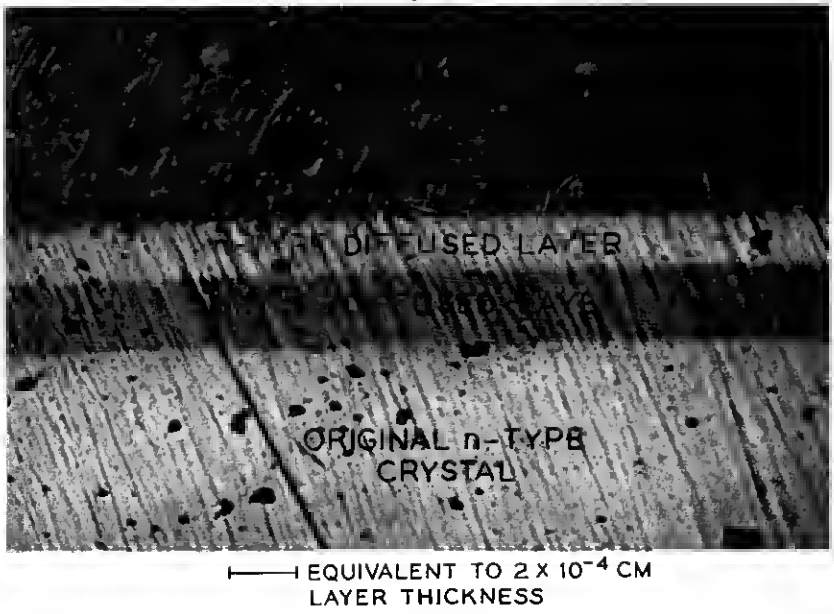

Fig. 1 — Angle section of a double diffused silicon wafer. The p-type center layer is approximately 2×10^{-4} cm thick.

niques employed in the fabrication of grown junction transistors, Fig. 2(b). However, a much simpler technique has been evolved. If the surface concentration of the donor diffusant is maintained below a certain critical value, it is possible to alloy an aluminum wire directly through the diffused n-type layer and thus make effective contact to the base layer, Fig. 2(c). Since the resistivity of the original silicon wafer is one to five ohm-cm, the aluminum will be rectifying to this region. It has been experimentally shown that if the surface concentration of the donor diffusant is less than the critical value mentioned above, the aluminum will also be rectifying to the diffused n-type region and the contact becomes merely an extension of the base layer. The n-layers produced by diffusing from elemental antimony are below the critical concentration and the direct aluminum alloying technique is feasible.

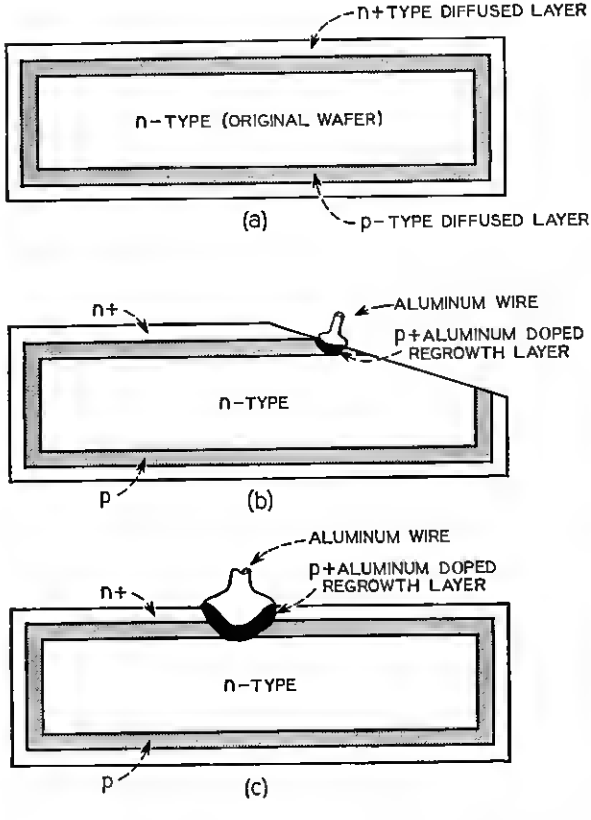

Fig. 2 — Schematic illustration of (a) double diffused n-p-n wafer, (b) angle section method of making base contact, and (c) direct alloying method of making base contact.

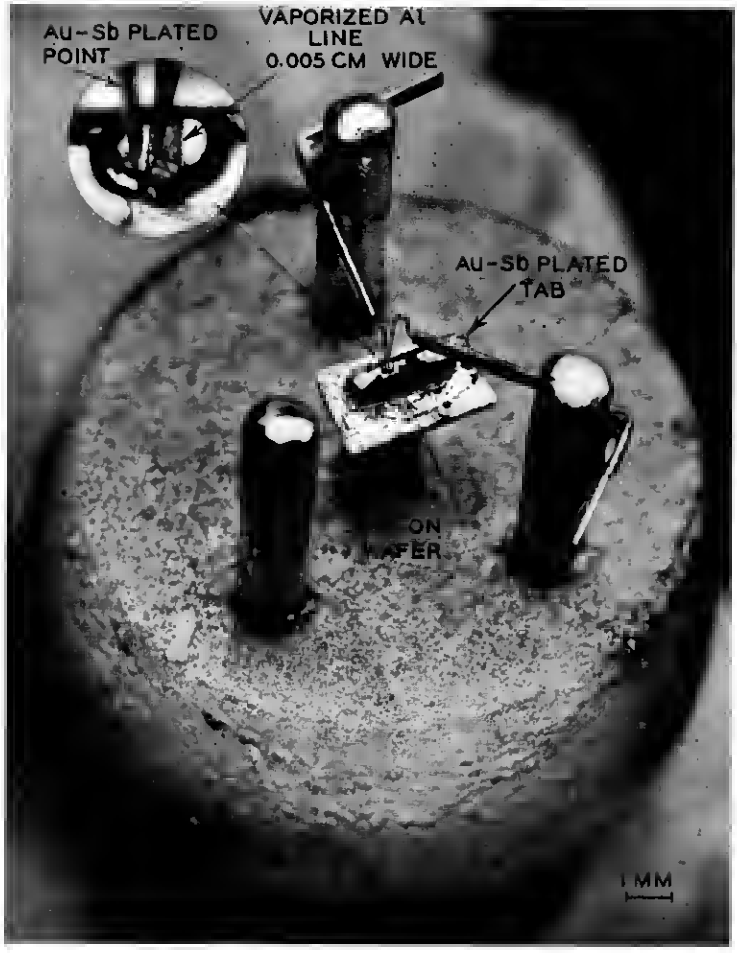

Fig. 3 — Mounted double diffused transistor.

Contact to the emitter layer is achieved by alloying a film of gold containing a small amount of antimony. Since this alloy will produce an n-type regrowth layer, it is only necessary to insure that the gold-antimony film does not alloy through the p-type base layer, thus shorting the emitter to the collector. This is controlled by limiting the amount of gold-antimony alloy which is available by using a thin evaporated film or by electroplating a thin film of gold-antimony alloy on an inert metal point and alloying this structure to the emitter layer.

Ohmic contact to the collector is produced by alloying the silicon wafer to an inert metal tab plated with a gold-antimony alloy.

Tbe transistors whose characteristics are reported in this paper were prepared from 3 ohm-cm n-type silicon using antimony and aluminum as the diffusants. The base contact was produced by evaporating aluminum through a mask so that a line approximately 0.005 × 0.015 cm in lateral dimensions and 100,000 Å thick was formed on the surface. This aluminum line was alloyed tbrougb the emitter layer in a subsequent operation. The wafer was then alloyed onto the plated kovar tab. A small area approximately 0.015 cm in diameter was masked around the line and the wafer was etched to remove the unwanted layers. The unit was then mounted in a header. Electrical contact to the collector was made by soldering to the kovar tab. Contact to the base was made with a tungsten point pressure contact to the alloyed aluminum. Contact to the emitter was made by bringing a gold-antimony plated tungsten point into pressure contact witb the emitter layer. The gold-antimony plate was then alloyed by passing a controlled electrical pulse between the plated point and the transistor collector lead. Fig. 3 is a photograph of a mounted unit.

2.0 ELECTRICAL CHARACTERISTICS

The frequency cutoffs of experimental double diffused silicon transistors fabricated as described above are an order of magnitude higher than the known cutoff frequencies of earlier silicon transistors. This is shown in Fig. 4 which gives the measured common base and common emitter current gains for one of these units as a function of frequency. The common base short-circuit current gain is seen to have a cutoff frequency of about 120 mc/sec.[2] The common emitter short-circuit current gain is shown on the same figure. The low-frequency current gain is better than thirty decibels and tbe cutoff frequency which is indicated by the frequency at which the gain is 3 db below its low-frequency value is 3 mc/sec. This is an exceptionally large common emitter bandwidth for a thirty db common emitter current gain and is of the same order of magnitude as that obtained with the highest frequency germanium transistors (e.q., p-n-i-p or tetrode) which had been made prior to the diffused base germanium transistor.[3]

[2] The increase in common base current gain above unity (indicated by current gain in decibels being positive) in the vicinity of 50 mc/sec is caused by a reactance gain error in the common base measurement. This error is caused by a combination of the emitter to ground parasitic capacitance and the positive reactance component of the transistor input impedance resulting from phase shift iu the alpha current gain.

[3] C. A. Lee, A High-Frequency Diffused Base Germanium Transistor, see page 23.

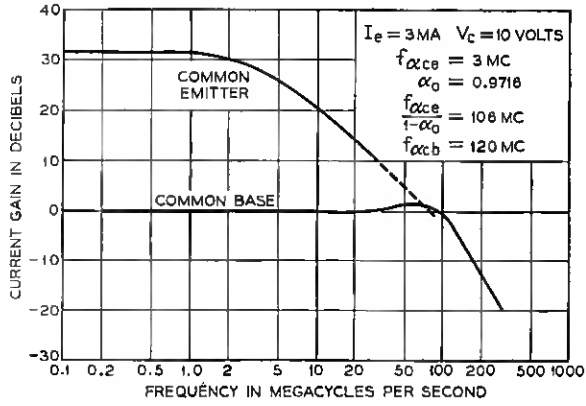

Fig. 4 — Short-circuit current gain of a double diffused silicon n-p-n transistor as a function of frequency in the common emitter and common base connections.

Fig. 5 shows a high-frequency lumped constant equivalent circuit for the double diffused silicon transistor whose current gain cutoff characteristic is shown in Fig. 4. External parasitic capacitances have been omitted from the circuit. The configuration is the conventional one for junction transistors with two exceptions. A series resistance r_e' has been added in the emitter circuit to account for contact resistance resulting from the fact that the present emitter point contacts are not perfectly ohmic. A second resistance r_c' has been added in the collector circuit to account for the ohmic resistance of the n-type silicon between the collector terminal and the effective collector junction. This resistance exists in all junction transistors but in larger area low frequency junction transistors its effect on alpha-cutoff is sufficiently small so that it has been ignored in equivalent circuits of these devices. The collector RC

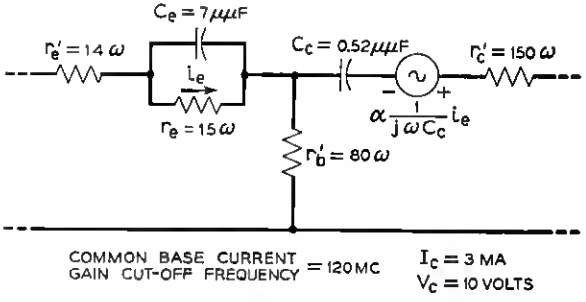

Fig. 5 — High-frequency lumped constant equivalent circuit for a double diffused silicon n-p-n transistor.

cutoff caused by the collector capacitance and the combined collector body resistance and base resistance is an order of magnitude higher than tbe measured alpha cutoff frequency and therefore is not too serious in impairing the very high-frequency performance of the transistor. This is due to the low capacitance of the collector junction which is seen to be approximately 0.5 mmf at 10 volts collector voltage. The base resistance of this transistor is less than 100 ohms which is quite low and compares very favorably with the best low frequency transistors reported previously.

The low-frequency characteristics of the double diffused silicon transistor are very similar to those of other junction transistors. This is illustrated in Fig. 6 where the static collector characteristics of one of these transistors are given. At zero emitter current the collector current is too small to be seen on the scale of this figure. The collector current

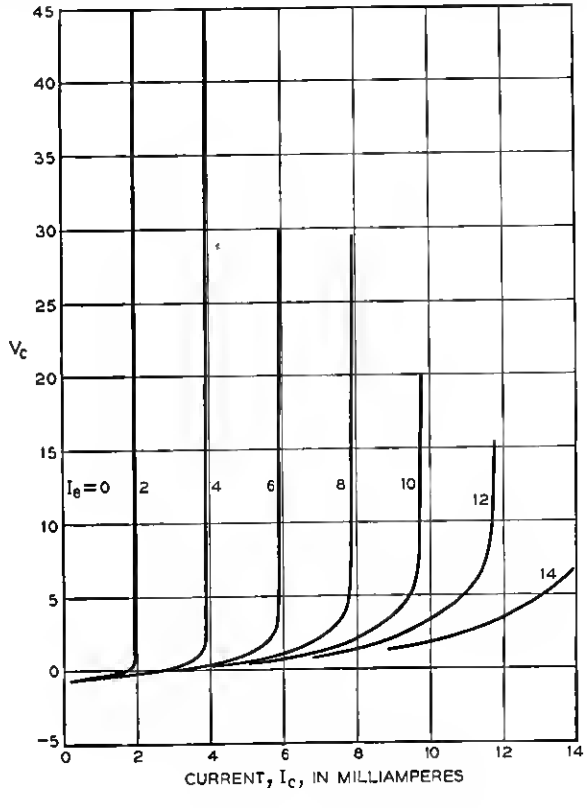

Fig. 6 — Collector characteristics of a double diffused silicon n-p-n transistor.

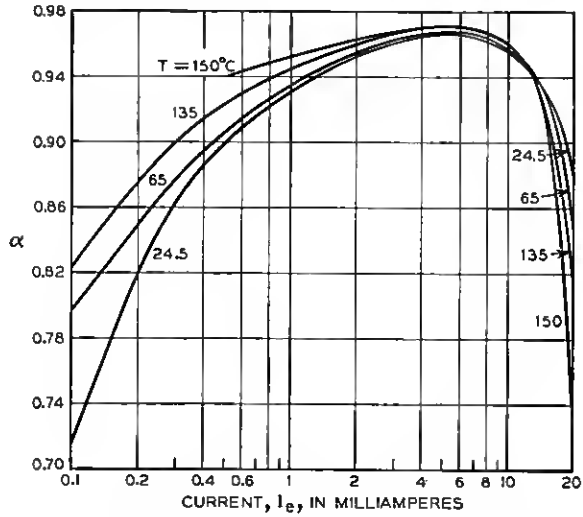

Fig. 7 — Alpha as a function of emitter current and temperature for a double diffused silicon n-p-n transistor.

under this condition does not truly saturate but collector junction resistance is very high. Collector junction resistances of 50 megohms at reverse biases of 50 volts are common.

The continuous power dissipation permissible with these units is also shown in Fig. 6. The figure shows dissipation of 200 milliwatts and the units have been operated at 400 milliwatts without damage. As illustrated in Fig. 3 no special provision has been made for power dissipation and it would appear from the performance obtained to date that powers of a few watts could be handled by these units with relatively minor provisions for heat dissipation. However, it can also be seen from Fig. 6 that at low collector voltages alpha decreases rapidly as the emitter current is increased. The transistor is, therefore, non-linear in this range of emitter currents and collector voltages. In many applications, this non-linearity may limit the operating range of the device to values below those which would be permissible from the point of view of continuous power dissipation.

Fig. 7 gives the magnitude of alpha as a function of emitter current for a fixed collector voltage of 10 volts and a number of ambient temperatures. These curves are presented to illustrate the stability of the parameters of the double diffused silicon transistor at increased ambient temperatures. Over the range from 1 to 15 milliamperes emitter current and 25°C to 150°C ambient temperature, alpha is seen to change only

by approximately 2 per cent. This amounts to only 150 parts per million change in alpha per degree centigrade change in ambient temperature.

The decrease in alpha at low emitter currents shown in Fig. 7 has been observed in every double diffused silicon transistor which has been made to date. Although this effect is not completely understood at present it could be caused by recombination centers in the base layer that can be saturated at high injection levels. Such saturation would result in an increase in effective lifetime and a corresponding increase in alpha. The large increase in alpha with temperature at low emitter currents is consistent with this proposal. It has also been observed that shining a strong light on the transistor will produce an appreciable increase in alpha at low emitter currents but has little effect at high emitter currents. A strong light would also be expected to saturate recombination centers which are active at low emitter currents and this behavior is also consistent with the above proposal.

3.0 DISCUSSION OF THE TRANSISTOR STRUCTURE

Although the low frequency electrical characteristics of the double diffused silicon transistor which are presented in Section 2 are quite similar to those usually obtained in junction transistors, the structure of the double diffused transistor is sufficiently different from that of the grown junction or alloy transistor that a discussion of some design principles is warranted. This section is devoted to a general discussion of the factors which determine the electrical characteristics of the transistors. In Section 4 the general ideas of Section 3 are applied in a more specialized fashion to the double diffused structure and a detailed calculation of electrical parameters is presented.

One essential difference between the double diffused transistor and grown junction or alloy transistors arises from the manner in which the impurities are distributed in the three active regions. In the ideal case of a double-doped grown junction transistor or an alloy transistor the concentration of impurities in a given region is essentially uniform and the transition from one conductivity type to another at the emitter and collector junctions is abrupt giving rise to step junctions. On the other hand in the double diffused structure the distribution of impurities is more closely described by the error function complement and the emitter and collector junctions are graded. These differences can have an appreciable influence on the electrical behavior of the transistors.

Fig. 8(a) shows the probable distribution of donor impurities, N_D, and acceptor impurities, N_A, in a double diffused n-p-n. Fig. 8(b) is a

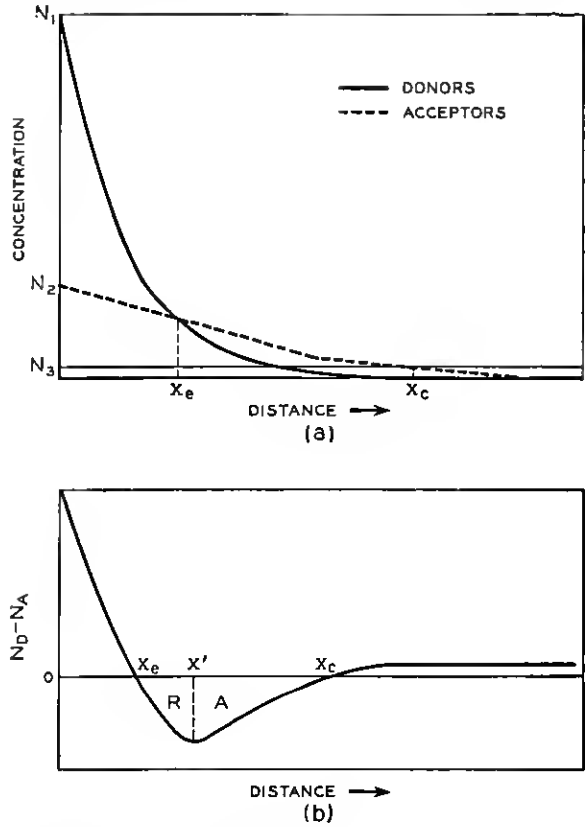

Fig. 8 — Diagrammatic representation of (a) donor and acceptor distributions and (b) uncompensated impurity distribution in a double diffused n-p-n transistor.

plot of $N_D - N_A$ which would result from the distribution in Fig. 8(a). Krömer[4] has shown that a nonuniform distribution of impurities in a semiconductor will produce electric fields which can influence the flow of electrons and holes. For example, in the base region the fields between the emitter junction, x_e, and the minimum in the $N_D - N_A$ curve, x', will retard the flow of electrons toward the collector while the fields between this minimum and the collector junction, x_c, will accelerate the flow of electrons toward the collector. These base layer fields will affect the transit time of minority carriers across the base and thus contribute

[4] H. Krömer, On Diffusion and Drift Transistor Theory I, II, III, Archiv. der Electr. Übertragung, **8**, pp. 223–228, pp. 363–369, pp. 499–504, 1954.

to the frequency response of the transistor. In addition the base resistance will be dependent on the distribution of both diffusants. These three factors are discussed in detail below.

Moll and Ross[5] have determined that the minority current, I_m, that will flow into the base region of a transistor if the base is doped in a nonuniform manner is given by

$$I_m = \frac{n_i^2 q D_m}{\int_b N(x)\, dx} e^{qV/kT} \tag{3.1}$$

where n_i is the carrier concentration in intrinsic material, q is the electronic charge, V is the applied voltage, D_m is the diffusion coefficient of the minority carriers, and the integral represents the total number of uncompensated impurities in the base. The primary assumptions in this derivation are (1) planar junctions, (2) no recombination in the base region, and (3) a boundary condition at the collector junction that the minority carrier density at this point equals zero. It is also assumed that the minority carrier concentration in the base region just adjacent to the emitter junction is equal to the equilibrium minority carrier density at this point multiplied by the Boltzman factor exp (qV/kT). It is of special interest to note that I_m depends only on the total number of uncompensated impurities in the base and not on the manner in which they are distributed.

In the double diffused transistor, it has been convenient from the point of ease of fabrication to make the emitter layer approximately the same thickness as the base layer. It has been observed that heating silicon to high temperatures degrades the lifetime of n- and p-type silicon in a similar manner.[6] Both base and emitter layers have experienced the same heat treatment and to a first approximation it can be assumed that the lifetime in the two regions will be essentially the same. Thus assumptions (1) and (2) should also apply to current flow from base to emitter. If we assume that the surface recombination velocity at the free surface of the emitter is infinite, then this imposes a boundary condition at this side of the emitter which under conditions of forward bias on the emitter is equivalent to assumption (3). Thus an equation of the form of (3.1) should also give the minority current flow from base to emitter. Since the emitter efficiency, γ, is given by

[5] J. L. Moll and I. M. Ross, The Dependence of Transistor Parameters on the Distribution of Base Layer Resistivity, Proc. I.R.E. in press.

[6] G. Bemski, private communication.

$$\gamma = \frac{I_m(\text{emitter to base})}{I_m(\text{emitter to base}) + I_m(\text{base to emitter})}$$

proper substitution of (3.1) will give the emitter efficiency of the double diffused n-p-n transistor,

$$\gamma = \frac{1}{1 + \frac{D_p}{D_n} \frac{\int_b (N_A - N_D)\, dx}{\int_e (N_D - N_A)\, dx}} \tag{3.2}$$

In (3.2), D_p is the diffusion coefficient of holes in the emitter, D_n is the diffusion coefficient of electrons in the base and the ratio of integrals is the ratio of total uncompensated doping in the base to that in the emitter.

A calculation of transit time is more difficult. Krömer[4] has studied the case of an aiding field which reduces transit time of minority carriers across the base region and thus increases frequency response. In the double diffused transistor the situation is more complex. Near the emitter side of the base region the field is retarding (Region R, see Fig. 8) and becomes aiding (Region A) only after the base region doping reaches a maximum. The case of retarding fields has been studied by Lee[3] and by Moll.[7] At present, the case for a base region containing both types of fields has not been solved. However, at the present state of knowledge some speculations about transit time can be made.

The two factors of primary importance are the magnitude of the built-in fields and the distance over which they extend. In the double diffused transistor, the widths of regions R and A are determined by the surface concentrations and diffusion coefficients of the diffusants. It can be shown by numerical computation[7] that if region R constitutes no more than 30–40 per cent of the entire base layer width, then the overall effect of the built-in fields will be to aid the transport of minority carriers and to produce a reduction in transit time. In addition the absolute magnitude of region R is important. If the point x' should occur within an effective Debye length from the emitter junction, i.e., if x' is located in the space charge region associated with the emitter junction, then the retarding fields can be neglected.

The base resistance can also be calculated from surface concentrations and diffusion coefficients of the impurities. This is done by considering the base layer as a conducting sheet and determining the sheet con-

[7] J. L. Moll, private communication.

ductivity from the total number of uncompensated impurities per square centimeter of sheet and the appropriate mobility weighted to account for impurity scattering.

4.0 CALCULATION OF DESIGN PARAMETERS

To calculate the parameters which determine emitter efficiency, transit time, and base resistance it is assumed that the distribution of uncompensated impurities is given by

$$N(x) = N_1 erfc \frac{x}{L_1} - N_2 erfc \frac{x}{L_2} + N_3 \tag{4.1}$$

where N_1 and N_2 are the surface concentrations of the emitter and base impurity diffusants respectively, L_1 and L_2 are their respective diffusion lengths, and N_3 is the original doping of the semiconductor into which the impurities are diffused. The impurity diffusion lengths are defined as

$$L_1 = 2\sqrt{D_1 t_1} \quad \text{and} \quad L_2 = 2\sqrt{D_2 t_2} \tag{4.2}$$

where the D's are the respective diffusion coefficients and the t's are the diffusion times.

Equation (4.1) can be reduced to

$$\Gamma(\xi) = \Gamma_1 \operatorname{erfc} \xi - \Gamma_2 \operatorname{erfc} \lambda\xi + 1 \tag{4.3}$$

where

$$\Gamma(\xi) = \frac{N(\xi)}{N_3}; \qquad \Gamma_1 = \frac{N_1}{N_3}; \qquad \Gamma_2 = \frac{N_2}{N_3}; \qquad \xi = \frac{x}{L_1}; \qquad \lambda = \frac{L_1}{L_2}$$

For cases of interest here, $\Gamma(\xi)$ will be zero at two points, α and β, and will have one minimum at ξ'. In the transistor structure the emitter junction occurs at $\xi = \alpha$ and the collector junction occurs at $\xi = \beta$. Thus the base width is determined by $\beta - \alpha$. The extent of aiding and retarding fields in the base is determined by ξ'. The integral of (4.3) from 0 to α, I_1, and from α to β, I_2, are the integrals of interest in (3.2) and thus determine emitter efficiency. In addition I_2 is the integral from which base resistance can be calculated.

The calculations which follow apply only for values of Γ_1/Γ_2 and Γ_2 greater than ten. Some of the simplifying assumptions which are made will not apply at lower values of these parameters where the distribution of both diffusants as well as the background doping affect the structure in all three regions of the device.

4.1 *Base Width*

From Fig. 8 and (4.3) it can be seen that for $\Gamma_2 \geqq 10$, α is essentially independent of Γ_2 and is primarily a function of Γ_1/Γ_2 and λ. Fig. 9 is a plot of α versus Γ_1/Γ_2 with λ as the parameter. The particular plot is for $\Gamma_2 = 10^4$. Although as stated α is essentially independent of Γ_2, at lower values of Γ_2, α may not exist for the larger values of λ, i.e., the p-layer does not form.

In the same manner, it can be seen that β is essentially independent of Γ_1/Γ_2 and is a function only of Γ_2 and λ. Fig. 10 is a plot of β versus Γ_2 with λ as a parameter. This plot is for $\Gamma_1/\Gamma_2 = 10$ and at larger Γ_1/Γ_2, β may not exist at large λ.

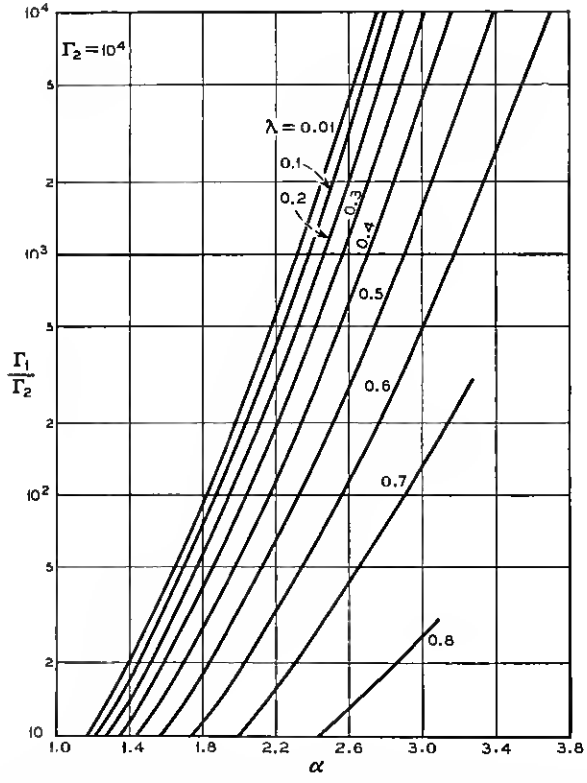

Fig. 9 — Emitter layer thickness (in reduced units) as a function of the ratio of the surface concentrations of the diffusing impurities (Γ_1/Γ_2) and the ratio of their diffusion lengths (λ).

Tbe base width

$$w = \beta - \alpha$$

can he obtained from Figs. 9 and 10. α, β and w can be converted to centimeters by multiplying hy the appropriate value of L_1.

4.2 *Emitter Efficiency*

With the limits α and β determined ahove, the integrals I_1 and I_2 can be calculated. Examination of the integrals shows that I_1 is closely proportional to Γ_1/Γ_2 and also to Γ_2. On the other hand I_2 is closely proportional to Γ_2 and essentially independent of Γ_1/Γ_2. Thus, the ratio of I_2/I_1 which determines γ depends primarily on Γ_1/Γ_2. Fig. 11 is a plot of the constant I_2/I_1 contours in the $\Gamma_1/\Gamma_2 - \lambda$ plane for I_2/I_1 in the range from -1.0 to -0.01. The graph is for $\Gamma_2 = 10^4$. Since from (3.2)

$$\gamma = \frac{1}{1 - \dfrac{D_p}{D_n}\dfrac{I_2}{I_1}} \tag{4.4}$$

for an n-p-n transistor, and assuming $D_p/D_n = \frac{1}{3}$ for silicon, then

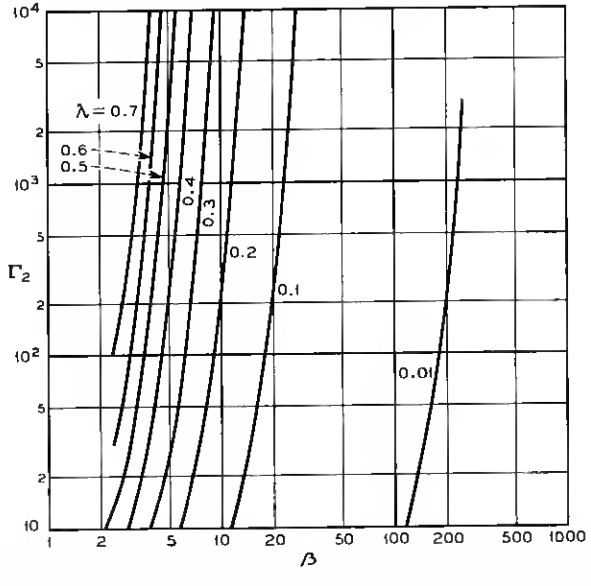

Fig. 10 — Collector junction depth (in reduced units) as a function of the surface concentration (in reduced units) of the diffusant which determines the conductivity type of the base layer (Γ_2) and the ratio of the diffusion lengths (λ) of the two diffusing impurities.

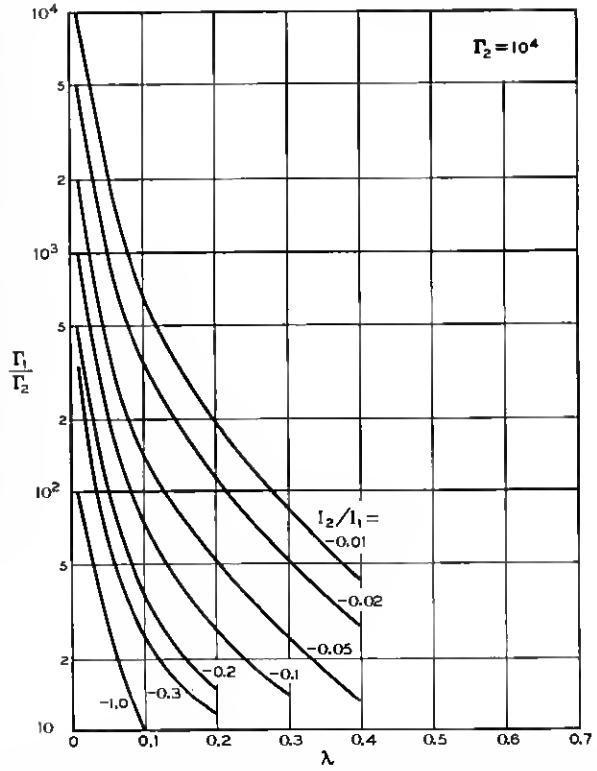

Fig. 11 — Dependence of emitter efficiency upon diffusant surface concentrations and diffusion lengths. The lines of constant I_2/I_1 are essentially lines of constant emitter efficiency. The ordinate is the ratio of surface concentrations of the two diffusants and the abscissa is the ratio of their diffusion lengths.

$I_2/I_1 = -1.0$ corresponds to a γ of 0.75 and $I_2/I_1 = -0.01$ corresponds to a γ of 0.997.

4.3 *Base Resistance*

It was indicated above that I_2 depends principally on Γ_2 and λ. Fig. 12 is a plot of the constant I_2 contours in the $\Gamma_2 - \lambda$ plane for I_2 in the range from -10^4 to -10. The graph is for $\Gamma_1/\Gamma_2 = 10$. The base layer sheet conductivity, g_b, can be calculated from these data as

$$g_b = -q\bar{\mu}I_2L_1N_3 \tag{4.5}$$

where q, L_1 and N_3 are as defined above and $\bar{\mu}$ is a mobility properly weighted to account for impurity scattering in the non-uniformly doped base region. The units of g_b are mhos per square.

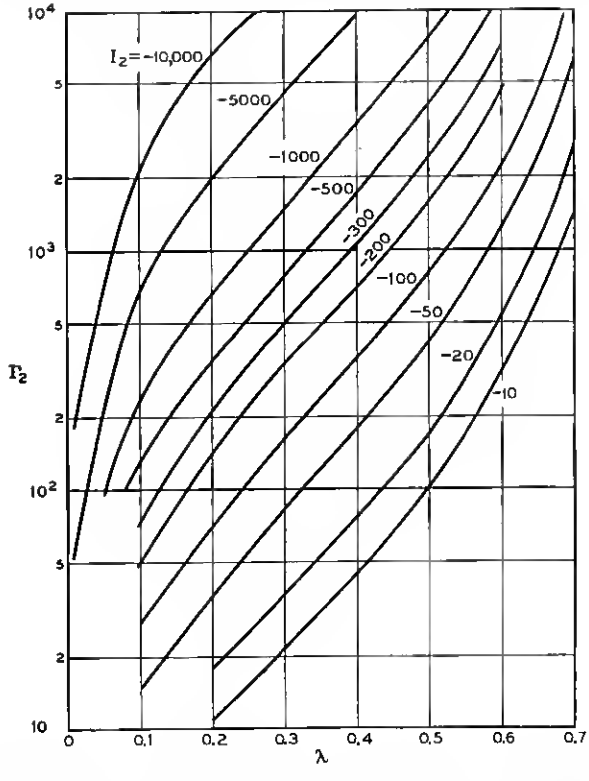

Fig. 12 — Dependence of base layer sheet conductivity on diffusant surface concentrations and diffusion lengths. The lines of constant I_2 are essentially lines of constant base sheet conductivity. The ordinate is the surface concentration (in reduced units) of the diffusant which determines the conductivity type of the base layer and the abscissa is the ratio of the diffusion lengths of the two diffusing impurities.

4.4 *Transit Time*

With a knowledge of where the minimum value, ξ', of (4.3) occurs, it is possible to calculate over what fraction of the base width the fields are retarding. The interesting quantity here is

$$\Delta R = \frac{\xi' - \alpha}{\beta - \alpha}$$

ξ' is a function of Γ_1/Γ_2 and λ and varies only very slowly with Γ_1/Γ_2. α is also a function of Γ_1/Γ_2 and λ and varies only slowly with Γ_1/Γ_2. The most rapidly changing part of ΔR is β which depends primarily on Γ_2 as noted above. Fig. 13 is a plot of the constant ΔR contours in the $\Gamma_2 - \lambda$ plane for values of ΔR in the range 0.1 to 0.3. This graph is

for data with $\Gamma_1/\Gamma_2 = 10$. As Γ_1/Γ_2 increases at constant Γ_2 and λ, ΔR decreases slightly. At $\Gamma_1/\Gamma_2 = 10^4$, the average change in ΔR is a decrease of about 25 per cent for constant Γ_2 and λ when $\Delta R \leqq 0.3$. The error is larger for values of ΔR greater than 0.3. It was noted above that when ΔR becomes greater than 0.3, the retarding fields become dominant. Therefore, this region is of slight interest in the design of a high frequency transistor.

4.5 *A Sample Design*

By superimposing Figs. 11, 12 and 13 the ranges of Γ_2, Γ_1/Γ_2 and λ which are consistent with desired values of γ, g_b and ΔR can be deter-

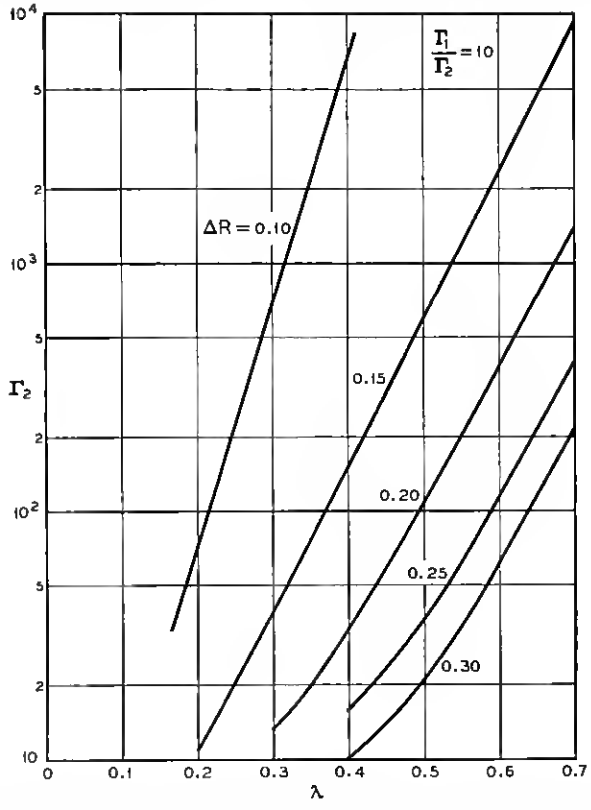

Fig. 13 — Dependence of the built-in field distribution on concentrations and diffusion lengths. The lines of constant ΔR indicate the fraction of the base layer thickness over which built-in fields are retarding. The ordinate is the surface concentration (in reduced units) of the diffusant which determines the conductivity type of the base layer and the abscissa is the ratio of the diffusion lengths of the two diffusing impurities.

mined by the area enclosed by the specified contour lines. It is also possible to compare the measured parameters of a specific device and observe how closely they agree with what is predicted from the estimated concentrations and diffusion coefficients. This is done below for the transistor described in Sections 1 and 2.

The comparison is complicated by the fact that the exact values of the surface concentrations and diffusion coefficients are not known precisely enough at present to permit an accurate evaluation of the design theory. However, the following values of concentrations and diffusion coefficients are thought to be realistic for this transistor.[1]

$$N_1 = 5 \times 10^{18} \qquad D_1 = 3 \times 10^{-12} \qquad t_1 = 5.7 \times 10^{3}$$

$$N_2 = 4 \times 10^{17} \qquad D_2 = 2.5 \times 10^{-11} \qquad t_2 = 1.2 \times 10^{3}$$

$$N_3 = 10^{15}$$

From these values it is seen that

$$\Gamma_1/\Gamma_2 = 12.5; \qquad \Gamma_2 = 400; \qquad \lambda = 0.6$$

From Fig. 9, $\alpha = 1.9$ and from Fig. 10, $\beta = 3.6$ and therefore $w = 1.7$. Measurement of the emitter and base layer dimensions showed that these layers were approximately the same thickness which was 3.8×10^{-4} cm. Thus the measured ratio of emitter width to base width of unity is in good agreement with the value of 1.1 predicted from the assumed concentrations and diffusion coefficients.

From Fig. 11, $I_2/I_1 \approx -0.01$. If this value is substituted into (4.4), $\gamma = 0.997$. This compares with a measured maximum alpha of 0.972.

From Fig. 12, $I_2 = -15$. Assuming an average hole mobility of 350 cm^2/volt. sec. and evaluating L_1 from the measured emitter thickness and the calculated α, (4.5) gives a value of $g_b = 1.7 \times 10^{-4}$ mhos per square. The geometry of the emitter and base contacts as shown in Fig. 3 makes it difficult to calculate the effective base resistance from the sheet conductivity even at very small emitter currents. In addition at the very high injection levels at which these transistors are operated the calculation of effective base resistance becomes very difficult. However, from the geometry it would be expected that the effective base resistance would be no greater than 0.1 of the sheet resistivity or 600 ohms. This is about seven times larger than the measured value of 80 ohms reported in Section 2.

From Fig. 13, ΔR is approximately 0.20. Thus there should be an overall aiding effect of the built-in fields. In addition the impurity gradient at the emitter junction is believed to be approximately $10^{21}/cm^4$ and the

space charge associated with this gradient will extend approximately 2×10^{-5} cm into the base region. The base thickness over which retarding fields extend is ΔR times the base width or 7.6×10^{-5} cm. Thus the first quarter of region R will be space charge and can be neglected.

The frequency cutoff from pure diffusion transit is given by

$$f_\alpha = \frac{2.43D}{2\pi W^2} \tag{4.6}$$

where W is the measured base layer thickness. Assuming $D = 25$ cm^2/sec for electrons in the base region, $f_\alpha = 67$ mc/sec. Since the measured cutoff was 120 mc/sec, the predicted aiding effect of the built-in field is evidently present.

These computations illustrate how the measured electrical parameters can be used to check the values of the surface concentrations and diffusion coefficients. Conversely knowledge of the concentrations and diffusion coefficients aid in the design of devices which will have prescribed electrical parameters. The agreement in the case of the transistor described above is not perfect and indicates errors in the proposed values of the concentrations and diffusion coefficients. However, it is sufficiently close to be encouraging and indicate the value of the calculations.

The discussion of design has been limited to a very few of the important parameters. Junction capacitances, emitter and collector resistances are among the other important characteristics which have been omitted here. Presumably all of these quantities can be calculated if the detailed structure of the device is known and the structure should be susceptible to the type of analysis used above. Another fact, which has been ignored, is that these transistors were operated at high injection levels and a low level analysis of electrical parameters was used. All of these other factors must be considered for a detailed understanding of the device. The object of this last section has been to indicate one path which the more detailed analysis might take.

5.0 CONCLUSIONS

By means of multiple diffusion, it has been possible to produce silicon transistors with alpha-cutoff above 100 mc/sec. Refinements of the described techniques offer the possibility of even higher frequency performance. These transistors show the other advantages expected from silicon such as low saturation currents and satisfactory operation at high temperatures.

The structure of the double diffused transistor is susceptible to design

analysis in a fashion similar to that which has been applied to other junction transistors. The non-uniform distribution of impurities produces significant electrical effects which can be controlled to enhance appreciably the high-frequency behavior of the devices.

The extreme control inherent in the use of diffusion to distribute impurities in a semiconductor structure suggests that this technique will become one of the most valuable in the fabrication of semiconductor devices.

ACKNOWLEDGEMENT

The authors are indebted to several people who contributed to the work described in this paper. In particular, the double diffused silicon from which the transistors were prepared was supplied by C. S. Fuller and J. A. Ditzenberger. The data on diffusion coefficients and concentrations were also obtained by them.

P. W. Foy and G. Kaminsky assisted in the fabrication and mounting of the transistors and J. M. Klein aided in the electrical characterization. The computations of the various solutions of the diffusion equation, (4.3), were performed by Francis Maier. In addition many valuable discussions with C. A. Lee, G. Weinreich, J. L. Moll, and G. C. Dacey helped formulate many of the ideas presented herein.